

# Distinct neural mechanisms of spatial attention and expectation guide perceptual inference in a multisensory world

Zuanazzi, Arianna; Noppeney, Uta

DOI:

[10.1523/JNEUROSCI.2873-18.2019](https://doi.org/10.1523/JNEUROSCI.2873-18.2019)

License:

None: All rights reserved

*Document Version*

Peer reviewed version

*Citation for published version (Harvard):*

Zuanazzi, A & Noppeney, U 2019, 'Distinct neural mechanisms of spatial attention and expectation guide perceptual inference in a multisensory world', *The Journal of Neuroscience*, vol. 39, no. 12, pp. 2301-2312. <https://doi.org/10.1523/JNEUROSCI.2873-18.2019>

[Link to publication on Research at Birmingham portal](#)

## **Publisher Rights Statement:**

Copyright © 2019 the authors

## **General rights**

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## **Take down policy**

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

---

**Research Articles: Behavioral/Cognitive**

**Distinct neural mechanisms of spatial attention and expectation guide perceptual inference in a multisensory world**

Arianna Zuanazzi<sup>1</sup> and Uta Noppeney<sup>1</sup>

<sup>1</sup>Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, B15 2TT Birmingham, UK

<https://doi.org/10.1523/JNEUROSCI.2873-18.2019>

Received: 7 November 2018

Revised: 8 January 2019

Accepted: 9 January 2019

Published: 18 January 2019

---

**Author contributions:** A.Z. and U.N. designed research; A.Z. performed research; A.Z. and U.N. analyzed data; A.Z. wrote the first draft of the paper; A.Z. and U.N. edited the paper; A.Z. and U.N. wrote the paper.

**Conflict of Interest:** The authors declare no competing financial interests.

This work was supported by the European Research Council (ERC-2012-StG\_20111109 multisens).

Corresponding author: Arianna Zuanazzi, Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, B15 2TT Birmingham, UK, [arianna.zuanazzi@gmail.com](mailto:arianna.zuanazzi@gmail.com)

**Cite as:** J. Neurosci 2019; 10.1523/JNEUROSCI.2873-18.2019

**Alerts:** Sign up at [www.jneurosci.org/alerts](http://www.jneurosci.org/alerts) to receive customized email alerts when the fully formatted version of this article is published.

**Distinct neural mechanisms of spatial attention and expectation guide perceptual inference  
in a multisensory world**

Abbreviated title: Multisensory spatial attention and expectation

Authors: Arianna Zuanazzi<sup>1</sup>, Uta Noppeney<sup>1</sup>

<sup>1</sup>Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and  
Cognitive Robotics Centre, University of Birmingham, B15 2TT Birmingham, UK

Corresponding author: Arianna Zuanazzi, Computational Cognitive Neuroimaging  
Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of  
Birmingham, B15 2TT Birmingham, UK, [arianna.zuanazzi@gmail.com](mailto:arianna.zuanazzi@gmail.com)

Competing interests: The authors declare no competing interests.

Funding: This work was supported by the European Research Council (ERC-2012-  
StG\_20111109 multisens).

23 **Abstract**

24 Spatial attention (i.e., task-relevance) and expectation (i.e., signal probability) are two critical  
25 top-down mechanisms guiding perceptual inference. Spatial attention prioritizes processing of  
26 information at task-relevant locations. Spatial expectations encode the statistical structure of the  
27 environment. An unresolved question is how the brain allocates attention and forms expectations  
28 in a multisensory environment, where task-relevance and signal probability over space can differ  
29 across sensory modalities.

30 We used functional magnetic resonance imaging in human participants (female and male) to  
31 investigate whether the brain encodes task-relevance and signal probability over space separately  
32 or interactively across sensory modalities. In a novel multisensory paradigm, we manipulated  
33 spatial attention and expectation selectively in audition and assessed their effects on behavioral  
34 and neural responses to auditory and visual stimuli.

35 Our results show that both auditory and visual stimuli increased activations in a right-lateralized  
36 fronto-parietal system, when they were presented at locations that were task-irrelevant in  
37 audition. Yet, only auditory stimuli increased activations in the medial prefrontal cortex when  
38 presented at expected locations and in audiovisual and fronto-parietal cortices signaling a  
39 prediction error when presented at unexpected locations.

40 This dissociation in multisensory generalization for attention and expectation effects shows that  
41 the brain controls attentional resources interactively across the senses but encodes the statistical  
42 structure of the environment as spatial expectations independently for each sensory system. Our  
43 results demonstrate that spatial attention and expectation engage partly overlapping neural  
44 systems via distinct mechanisms to guide perceptual inference in a multisensory world.

45

46 **Significance statement**

47 In our natural environment the brain is exposed to a constant influx of signals through all our  
48 senses. How does the brain allocate attention and forms spatial expectations in this multisensory  
49 environment? Because observers need to respond to stimuli irrespective of their sensory  
50 modality, they may allocate attentional resources and encode the probability of events jointly  
51 across the senses.

52 This psychophysics and neuroimaging study shows that the brain controls attentional resources  
53 interactively across the senses via a fronto-parietal system but encodes the statistical structure of  
54 the environment independently for each sense in sensory and fronto-parietal areas. Thus, spatial  
55 attention and expectation engage partly overlapping neural systems via distinct mechanisms to  
56 guide perceptual inference in a multisensory world.

57

58

59

60

## 61 Introduction

62 Spatial attention (i.e., task-relevance) and expectation (i.e., signal probability) are two critical  
63 top-down mechanisms that guide perceptual inference. Spatial attention prioritizes signal  
64 processing at locations that are relevant for the observer's goals. Spatial expectations encode the  
65 event probability over space, i.e., the statistical structure of the environment (Summerfield and  
66 Egner, 2009).

67 Behaviorally, both spatial attention and expectation typically facilitate perception leading to  
68 faster and more accurate responses for stimuli presented at attended and/or expected locations  
69 (Posner et al., 1980; Downing, 1988; Doherty et al., 2005; Geng and Behrmann, 2002, 2005;  
70 Carrasco, 2011). At the neural level, spatial attention is thought to increase stimulus evoked  
71 responses at task-relevant locations (Tootell et al., 1998; Brefczynski and DeYoe, 1999; Bressler  
72 et al., 2013), whereas expectations often reduce stimulus-evoked responses (Alink et al., 2010;  
73 Summerfield et al., 2008; Kok et al., 2012a; though see: Kok et al., 2012b). Importantly, spatial  
74 attention and expectations are intimately related (Zuanazzi and Noppeney, 2018). In many  
75 situations observers will allocate attentional resources to locations where events are likely to  
76 occur (Summerfield and Egner, 2009). Likewise, the majority of previous paradigms, most  
77 prominently the classical Posner paradigm (Posner, 1980), manipulated observer's endogenous  
78 spatial attention via probabilistic cues that indicate where a task-relevant target is likely to occur.  
79 Only recently has unisensory research attempted to dissociate attention and expectation  
80 (Shulman et al., 2009; Doricchi et al., 2010; Kok et al., 2012b; Aukstulewicz and Friston,  
81 2015). Recent accounts of predictive coding suggest that attention may increase the precision of  
82 prediction errors that are elicited when expectations are violated (Feldman and Friston, 2010;  
83 Aukstulewicz and Friston, 2015).

84 Crucially, in our natural environment the brain is exposed to a constant influx of signals  
85 furnished by all our senses. This raises the critical question of how the brain allocates spatial  
86 attention and forms spatial expectations in a multisensory environment. Because observers need  
87 to respond to stimuli irrespective of the sense by which they are perceived, they may allocate  
88 attentional resources interactively across the senses and form an ‘amodal map’ that encodes the  
89 probability of events. In line with this conjecture, parietal cortices have previously been shown to  
90 integrate audiovisual signals weighted by their bottom-up sensory reliabilities and top-down  
91 task-relevance into audiovisual spatial priority maps (Rohe and Noppeney, 2015, 2016).  
92 Likewise, attentional resources were shown to be allocated interactively across the senses. Shifts  
93 in spatial attention that were endogenously or exogenously induced in one sensory modality  
94 affected stimulus processing in other sensory systems (Spence and Driver, 1996, 1997; Eimer  
95 and Schröger, 1998; Eimer, 1999; McDonald et al., 2000; Spence et al., 2000; Ward et al., 2000).  
96 Irrespective of stimulus modality, reorienting of spatial attention was associated with activations  
97 in ventral and to some extent dorsal fronto-parietal cortices (Corbetta and Shulman, 2002; Wu et  
98 al., 2007; Corbetta et al., 2008; Santangelo et al., 2008; Krumbholz et al., 2009; Macaluso, 2010;  
99 Santangelo and Macaluso, 2012).

100 Less is known about how the brain forms spatial expectations across sensory modalities  
101 (Stekelenburg and Vroomen, 2012). Because information is initially gathered by distinct sensory  
102 organs and enters the brain via parallel pathways, each sensory system may initially encode the  
103 probability of signals selectively for its preferred sensory modality. These modality-specific  
104 spatial expectations may be reinforced particularly in environments where auditory and visual  
105 signals arise from separate sources such as in experiments that present auditory or visual signals  
106 independently (Spence and Driver, 1996).

107 The current study investigated how the brain allocates spatial attention and forms spatial  
 108 expectations across the senses. Further, we assessed whether spatial attention and expectation  
 109 rely on distinct or common neural systems and guide perceptual inference via additive or  
 110 interactive mechanisms. Combining fMRI and a novel multisensory paradigm we orthogonally  
 111 manipulated spatial attention (i.e., task-relevance) and expectation (i.e., spatial signal  
 112 probability) selectively in audition and assessed their effects on observers' behavioral and neural  
 113 responses in audition and vision. We expected attentional resources to be interactively allocated  
 114 across sensory modalities (Eimer and Schröger, 1998; Macaluso et al., 2002). By contrast, given  
 115 the hierarchical organization of multisensory integration, spatial expectations and prediction  
 116 errors for unexpected stimuli may be modality-specific in early sensory cortices but shared  
 117 across the senses in parietal cortices (Rohe and Noppeney, 2015, 2016, 2018).

118

## 119 **Materials and Methods**

### 120 *Participants*

121 Thirty-one healthy volunteers (8 males; mean age: 21.4 years; range: 18–27 years) participated in  
 122 the psychophysics experiment. All participants had normal or corrected to normal vision,  
 123 reported normal hearing and had no history of neurological or psychiatric illness. All participants  
 124 were right-handed, according to the Edinburgh Handedness Inventory (Oldfield, 1971) (mean  
 125 laterality index: 84; range: 60–100). A subgroup of 22 participants (5 males; mean age: 21.2  
 126 years; range: 18–27 years) was selected to take part in the fMRI experiment (see Inclusion  
 127 criteria). Data collection was terminated when 22 participants had undergone the fMRI study.  
 128 This sample size was determined based on Thirion et al. (2007). All participants provided written  
 129 informed consent, as approved by the local ethics committee of the University of Birmingham



130 (Science, Technology, Mathematics and Engineering (STEM) Ethical Review Committee) and  
 131 the experiment was conducted in accordance with these guidelines and regulations.

132

### 133 *Inclusion criteria*

134 A subgroup of 22 participants who had taken part in the psychophysics experiment was selected  
 135 to take part in the fMRI experiment. Inclusion criteria were participants' accuracy and fixation  
 136 performance in the psychophysics experiment. Only participants who in the psychophysics  
 137 experiment produced less than 20 saccades averaged across blocks and showed overall accuracy  
 138 > 95% (calculated as the percentage of hits + correct rejections, pooling over auditory and visual  
 139 stimuli) were selected for the fMRI experiment.

140

### 141 *Stimuli*

142 Auditory spatialized stimuli (100 ms duration) were created by convolving a burst of white noise  
 143 (with 5 ms onset and offset ramps) with spatially specific head-related transfer functions  
 144 (HRTFs) based on the KEMAR dummy head of the MIT Media Lab  
 145 (<http://sound.media.mit.edu/resources/KEMAR.html>, Gardner and Martin, 1995).

146 Visual stimuli (i.e., the so-called 'flashes') were white discs (100 ms duration; radius: 0.88°  
 147 visual angle, luminance: 165 cd/m<sup>2</sup>) presented on a gray background (luminance: 78 cd/m<sup>2</sup>).

148 Both auditory and visual stimuli were presented at  $\pm 10^\circ$  visual angle along the azimuth (0° visual  
 149 angle for elevation). A fixation cross was presented in the center of the screen throughout the  
 150 entire experiment.

151

### 152 *Experimental design*

153 In both the psychophysics and the fMRI experiment, we orthogonally manipulated spatial  
154 attention (i.e., task-relevance or response requirement) and expectation (i.e., stimulus  
155 probability) across the two hemifields selectively in audition and evaluated their effects on  
156 observers' neural and behavioral responses to auditory and visual signals. Thus, the  $2 \times 2 \times 2 \times 2$   
157 design manipulated auditory spatial attention (left vs right hemifield), auditory spatial  
158 expectation (left vs right hemifield), stimulus location (left vs right hemifield) and stimulus  
159 modality (auditory vs visual, see Fig. 1A). For the behavioral and fMRI data analysis we pooled  
160 over stimulus locations (left/right) leading to a  $2$  (attended vs unattended)  $\times 2$  (expected vs  
161 unexpected)  $\times 2$  (auditory vs visual stimulus modality) factorial design. Across days, auditory  
162 spatial expectation was manipulated as spatial signal probability, i.e., the probability for auditory  
163 stimuli to be presented in the left or right hemifield. Both the psychophysics and fMRI  
164 experiments were preceded by training runs, in which the spatial probability ratio of auditory  
165 targets was set to 9/1 for the expected/unexpected hemifields to boost the implicit learning of  
166 auditory spatial signal probability. In the psychophysics and fMRI experiments the auditory  
167 stimuli were presented with a ratio of 4/1 in the expected/unexpected hemifields. Observers were  
168 not explicitly informed about those probabilities. Auditory spatial attention was manipulated as  
169 'task-relevance', i.e., the requirement to respond to an auditory target in the left vs right  
170 hemifield. Critically, spatial attention and expectation were manipulated only in audition but not  
171 in vision. Participants needed to respond to all visual targets which were presented in either  
172 spatial hemifield with equal probability (i.e., 1/1 in the expected/unexpected hemifields, see Fig.  
173 1A and 1B). Throughout the entire experiment a central fixation cross coded in color whether  
174 participants should attend and respond to sounds in either their left or right hemifield. The  
175 mapping between color and task-relevant hemifield was counterbalanced across participants.

176 *Spatial signal, general response and spatially selective response probability*

177 Our experiment orthogonally manipulated spatial attention as task-relevance and expectation as  
178 spatial signal probability selectively in audition. The attentional manipulation is therefore  
179 operationally linked with response requirement over space. Further, attention as response  
180 requirement and expectation as signal probability are intimately linked by co-determining  
181 general (i.e., the probability that the observer needs to make a response irrespective of the  
182 hemifield in which the signal is presented) and spatially selective (i.e., the probability that the  
183 observer needs to make a response conditioned on that the signal is presented in a particular  
184 hemifield) response probabilities.

185 As shown in figures 1A and 1B, the general response probability is greater in block type 1,  
186 where attention and expectation are directed to the same hemifields, than in block type 2, where  
187 attention and expectation are directed to different hemifields. Put differently, greater demands  
188 are placed on response inhibition in block type 2 where the hemifield with the more frequent  
189 auditory stimuli is task-irrelevant (i.e., a response needs to be inhibited).

190 Likewise, the spatially selective response probability is co-determined by both attention and  
191 expectation. Observers need to respond to both auditory and visual stimuli in the attended  
192 hemifield, so that the response probability in the attended hemifield is always equal to one. By  
193 contrast, in the unattended hemifield observers need to respond only to the visual stimuli. Hence,  
194 in the unattended hemifield the response probability also depends on the frequency of the  
195 auditory stimuli and hence on expectation. In the unattended hemifield the response probability  
196 is thus smaller and hence response inhibition greater when the task-irrelevant auditory stimuli are  
197 more frequent.

198 Importantly, in our paradigm general and spatially selective response probabilities would predict  
 199 an interaction between attention and expectation that is common to auditory and visual stimuli.  
 200 Conversely, main effects of attention and expectation cannot be explained by differences in  
 201 response probability.

202

### 203 *Experimental procedure*

204 The current study included two experiments: (i) a psychophysics experiment conducted across  
 205 two days (i.e., auditory spatial expectation was manipulated between the two days) and (ii) an  
 206 fMRI experiment conducted across other two days (i.e., auditory spatial expectation was  
 207 manipulated between the two days). The psychophysics experiment was conducted prior to the  
 208 fMRI experiment. On each day, the psychophysics and the fMRI experimental runs were  
 209 preceded by two training runs (see Experimental design).

210 Each experimental run (duration: ~8min/run) included 10 attention blocks with 20 trials each,  
 211 interleaved with 6 s fixation baseline periods. As a result of our balanced factorial design, blocks  
 212 were of two types: in block type 1, spatial attention and expectation were congruent (i.e., spatial  
 213 attention was directed to the hemifield with higher auditory target frequency); in block type 2,  
 214 spatial attention and expectation were incongruent (i.e., attention was directed to the hemifield  
 215 with lower auditory target frequency) (Fig. 1B). Thus, both psychophysics and fMRI  
 216 experiments included 2000 trials = 20 trials x 10 blocks (attention manipulation: 5 blocks of type  
 217 1 and 5 blocks of type 2) x 5 experimental runs x 2 days (expectation manipulation) in total.  
 218 Therefore, each block type included 400 auditory stimuli for the expected hemifield (pooled over  
 219 left and right) and 100 auditory stimuli for the unexpected hemifield (pooled over left and right).  
 220 Each block type also included 250 visual stimuli for the expected hemifield and 250 visual

221 stimuli for the unexpected hemifield (pooled over left and right). For further details see Fig. 1B  
 222 which shows the absolute number of trials for each condition and block type and their response  
 223 requirement for the psychophysics and the fMRI experiment.

224 The order of ‘expectation’ days (i.e., left vs right) and the color (i.e., pink or green) of the  
 225 fixation cross (i.e., attention instruction) were counterbalanced across participants, the order of  
 226 attention blocks was counterbalanced within and across participants and the order of stimulus  
 227 location and stimulus modality were pseudo-randomized within each participant. Brief breaks  
 228 were included after every run to provide feedback to participants about their performance  
 229 accuracy (averaged across all conditions) in the target detection task. In the psychophysics  
 230 experiment participants’ fixation performance was monitored via eye tracking, and participants  
 231 were provided with feedback about their eye-movements (i.e., fixation maintenance) during the  
 232 breaks. For the psychophysics experiment, mean group number of saccades was  $22.9 \pm 5.2$   
 233 [across subjects mean  $\pm$  SEM] and mean group accuracy was  $97\% \pm 0.2\%$  [across subjects mean  
 234  $\pm$  SEM] for the psychophysics experiment and  $97\% \pm 0.5\%$  [across subjects mean  $\pm$  SEM] for  
 235 the fMRI experiment.

236 Each trial (SOA: 2200 ms) included three periods (see Fig. 1C): i. the fixation cross alone (700  
 237 ms duration), ii. the brief flash or sound (stimulus duration: 100 ms) and iii. the fixation cross  
 238 alone, i.e., response window (1400 ms). Participants responded to the auditory targets in the  
 239 attended hemifield and to all visual targets via key press with their right index finger (i.e., the  
 240 same response for all auditory and visual targets) as fast and accurately as possible. They fixated  
 241 the cross in the center of the screen which was presented throughout the entire experiment.

242 On each day, participants were first familiarized with the stimuli in brief practice runs (with  
 243 equal spatial signal probability) to train them on target detection performance and, only in the

244 psychophysics experiment, also on fixation (i.e., a warning signal was shown when the disparity  
 245 between the central fixation cross and the eye-data samples exceeded 2.5 degrees).  
 246 After the final fMRI day, participants indicated in a questionnaire whether they thought the  
 247 sound or the flash was presented more frequently in one of the two spatial hemifields. Eighteen  
 248 out of the total 22 participants correctly reported that the auditory stimuli were more frequent in  
 249 one hemifield and 20 out of 22 participants reported the visual stimuli to be equally frequent  
 250 across the two hemifields, suggesting that most participants were aware of the manipulation of  
 251 signal probability.

252

253 Figure 1 approximately here

254

## 255 ***Experimental setup***

256 *Psychophysics experiment.* The psychophysics experiment (training and experimental runs) was  
 257 conducted in a darkened room. Participants rested their chin on a chinrest with the height held  
 258 constant across all participants. Auditory stimuli were presented at approximately 72 dB SPL,  
 259 via HD 280 PRO headphones (Sennheiser, Germany). To mimic the scanner environment, the  
 260 scanner noise was reproduced for the whole duration of the experiment at approximately 80 dB  
 261 SPL via external loudspeakers. Visual stimuli were displayed on a gamma-corrected LCD  
 262 monitor (2560 x 1600 resolution, 60 Hz refresh rate, 30" Dell UltraSharp U3014, USA), at a  
 263 viewing distance of approximately 50 cm from the participant's eyes. Stimuli were presented  
 264 using Psychtoolbox version 3 (Brainard, 1997; [www.psychtoolbox.org](http://www.psychtoolbox.org), RRID: SCR\_002881),  
 265 running under MATLAB R2014a (Mathworks Inc., Natick, MA, USA, RRID: SCR\_001622) on  
 266 a Windows machine. Participants responded to all targets with their right index finger and

267 responses were recorded via one key of a small keypad (Targus, USA). Throughout the study,  
 268 participants' eye-movements and fixations were monitored using Tobii Eyex eyetracking system  
 269 (Tobii, Sweden).

270 *fMRI experiment.* During the training runs, participants lay in a mock scanner, which mimicked  
 271 all features of the MRI scanner. The scanner noise was reproduced at approximately 80 dB SPL  
 272 via external loudspeakers. During the experimental runs, participants lay in the MRI scanner.  
 273 Auditory stimuli were presented at approximately 72 dB SPL using MR-compatible headphones  
 274 (MR Confon). Visual stimuli were back-projected onto a Plexiglas screen using a BARCO  
 275 projector (F35). Participants viewed the screen through a mirror mounted on the MR head coil at  
 276 a viewing distance of approximately 68 cm. Stimuli were presented using Psychtoolbox version  
 277 3 (Brainard, 1997; [www.psychtoolbox.org](http://www.psychtoolbox.org), RRID: SCR\_002881), running under MATLAB  
 278 R2014a (Mathworks Inc., Natick, MA, USA, RRID: SCR\_001622) on a MacBook Pro machine.  
 279 Participants responded to all targets with their right index finger and responses were recorded via  
 280 an MR-compatible keypad (NATA).

281

#### 282 *fMRI data acquisition*

283 A 3T Philips MRI scanner with 32 channel head coil was used to acquire both T1-weighted  
 284 anatomical images (TR = 8.4 ms, TE = 3.8 ms, flip angle = 8°, FOV = 288 mm x 232 mm, image  
 285 matrix = 288 x 232, 175 sagittal slices acquired in ascending direction, voxel size = 1 mm x  
 286 1mm x 1 mm) and T2\*-weighted axial echoplanar images (EPI) with bold oxygenation level-  
 287 dependent (BOLD) contrast (TR = 2600 ms, TE = 40 ms, flip angle = 85°, FOV = 240 mm x 240  
 288 mm, image matrix 80 x 80, 38 transversal slices acquired in ascending direction, voxel size = 3 x  
 289 3 x 3 mm). For each participant, an overall of 196 volumes x 5 experimental runs x 2 days =

1960 volumes were acquired. The anatomical image volume was acquired at the end of the experiment.

### *Statistical analysis*

#### *Behavioral data analysis - psychophysics and fMRI experiments*

For the behavioral analysis of the psychophysics experiment, we excluded trials where participants did not successfully fixate the central cross based on a dispersion criterion (i.e., distance of fixation from subject's median of fixation (as defined in calibration trials) > 1.3 degrees for three subsequent samples, Blignaut, 2009). Percentage [across subjects mean  $\pm$  SEM] of trials excluded for auditory stimuli: 1.4%  $\pm$  0.4%; for visual stimuli: 1.3%  $\pm$  0.4%. The response time analysis was limited to correct trials and response times within the range of participant- and condition-specific mean  $\pm$  two SD and < 1400 ms (i.e., within the response window).

For auditory targets in the attended hemifield, median response times for each subject were entered into a two-sided paired t-test with auditory spatial expectation (expected vs unexpected stimulus) as factor.

For visual targets, median response times for each subject were entered into a 2 (auditory spatial attention: attended vs unattended stimulus)  $\times$  2 (auditory spatial expectation: expected vs unexpected stimulus) repeated measures analysis of variance (ANOVA).

Unless otherwise indicated, we only report effects that are significant at  $p < 0.05$ .

#### *fMRI data analysis*



312 The functional MRI data were analyzed with statistical parametric mapping (SPM12; Wellcome  
 313 Department of Imaging Neuroscience, London; [www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm), Friston et al., 1995).  
 314 Scans from each subject were realigned using the first as a reference, unwarped, slice-time  
 315 corrected and spatially normalized into MNI standard space using parameters from segmentation  
 316 of the T1 structural image (Ashburner and Friston, 2005), resampled to a spatial resolution of 2 x  
 317 2 x 2 mm<sup>3</sup> and spatially smoothed with a Gaussian kernel of 8 mm full-width at half-maximum.  
 318 The time series of all voxels were high-pass filtered to 1/128 Hz.  
 319 The fMRI experiment was modeled in a mixed block/event-related fashion with regressors  
 320 entered into the design matrix after convolving each event-related unit impulse with a canonical  
 321 hemodynamic response function and its first temporal derivative. In addition to modeling the 16  
 322 conditions in our 2 (stimulus modality: auditory vs visual) x 2 (auditory spatial attention: left vs  
 323 right hemifield) x 2 (auditory spatial expectation: left vs right hemifield) x 2 (stimulus location:  
 324 left vs right hemifield) factorial design, the statistical model included the onsets of the attention  
 325 cue (i.e., auditory attention to the left hemifield, auditory attention to the right hemifield) as a  
 326 separate regressor. Nuisance covariates included the realignment parameters to account for  
 327 residual motion artifacts.  
 328 Condition-specific effects for each subject were estimated according to the general linear model  
 329 and passed to a second-level analysis as contrasts. This involved creating 16 contrast images  
 330 (i.e., each of the 16 conditions relative to fixation, summed over the 10 runs) for each subject and  
 331 entering them into a second-level ANOVA. Inferences were made at the second level to allow a  
 332 random-effects analysis and inferences at the population level (Friston et al., 1995).  
 333 At the random effects or group level, we pooled over stimulus locations (left/right) and,  
 334 separately for each sensory modality, we tested for (i) the main effect of spatial attention (i.e.,

335 attended > unattended auditory stimuli and vice versa, attended > unattended visual stimuli and  
 336 vice versa) and (ii) the main effect of spatial expectation (i.e., expected > unexpected auditory  
 337 stimuli and vice versa, expected > unexpected visual stimuli and vice versa).

338 To assess whether these effects of spatial attention and expectation rely on amodal or modality-  
 339 specific systems, we investigated (i) whether the effects of attention and expectation are common  
 340 for audition and vision (i.e., a logical “AND” conjunction over stimulus modalities) or (ii)  
 341 whether the effects differ between audition and vision (i.e., the interaction between attention and  
 342 stimulus modality and the interaction between expectation and stimulus modality).

343 Finally, we investigated whether spatial attention and expectation effects were additive or  
 344 interactive. Separately for each stimulus modality, we tested for (i) the effects that are common  
 345 for attention and expectation (i.e., a logical “AND” conjunction over each attention and  
 346 expectation main effects, i.e., additive effects) and (ii) the interaction between attention and  
 347 expectation.

348 Unless otherwise stated, we report activations at  $p < 0.05$  at the cluster level corrected for  
 349 multiple comparisons within the entire brain using an auxiliary (uncorrected) voxel threshold of  
 350  $p < 0.001$ .

351

### 352 ***ROI analysis***

353 Based on our a priori hypothesis that spatial attention and expectation influence activations in  
 354 primary sensory cortices, we tested for the effects of auditory spatial attention and expectation  
 355 selectively within the primary auditory cortex and primary visual cortex. These areas of interests  
 356 were defined using bilateral ROI maps from SPM Anatomy Toolbox (version 2.2b, Eickhoff et  
 357 al., 2005). The anatomical mask for the primary auditory cortex encompassed 890 voxels in the

358 bilateral cytoarchitectonic maps TE 1.0, TE 1.1 and TE 1.2. The anatomical mask for the primary  
 359 visual cortex encompassed 2936 voxels in the bilateral cytoarchitectonic maps hOC1. We  
 360 extracted parameter estimates from each ROI, for each of the 16 conditions relative to fixation  
 361 and for each subject and entered them into a 2 (auditory spatial attention: attended vs unattended  
 362 stimulus) x 2 (auditory spatial expectation: expected vs unexpected stimulus) repeated measures  
 363 ANOVA, separately for each stimulus modality (pooling over stimulus locations).  
 364 To sensitize our analysis to attentional modulation of evoked responses, we re-performed this  
 365 repeated measures ANOVA separately for (1) the ‘unilateral ROIs’ ipsilateral to the stimulus  
 366 location and (2) the ‘unilateral ROIs’ contralateral to the stimulus location. Practically, this  
 367 involved normalization to a symmetric MNI standard template (created by averaging the  
 368 standard MNI template with its flipped version, Didelot et al., 2010) and (1) pooling over  
 369 activations in the left ROI (for stimuli in the left hemifield) and the right ROI (for stimuli in the  
 370 right hemifield) (i.e., ipsilateral ROIs) and (2) pooling over activations in the left ROI (for  
 371 stimuli in the right hemifield) and the right ROI (for stimuli in the left hemifield) (i.e.,  
 372 contralateral ROIs), for the corresponding conditions in our 2 (attention) x 2 (expectation) x 2  
 373 (stimulus modality) design (for similar analyses, see Lipschutz et al., 2002; Macaluso and Patria,  
 374 2007). Because the results of these two ‘flipped’ analyses that separately tested for the effects of  
 375 attention and expectation on ipsilateral and contralateral stimuli were comparable (with small  
 376 deviations in p-values) to our main ROI analysis (because these effects were anyhow bilateral),  
 377 we do not report these results.

378

## 379 **Results**

380 In the following, we report (1) the behavioral results of the psychophysics and the fMRI  
 381 experiment and (2) the imaging results of the fMRI experiment.

382

### 383 ***Behavioral results – psychophysics and fMRI experiments***

384 In a target detection task, participants responded to auditory targets presented in their attended  
 385 hemifield (i.e., auditory attention manipulation) and to all visual targets (Fig.1A, 1B and 1C).

386 For both psychophysics and fMRI experiments, the two-sided paired-sample t-tests on response  
 387 times for auditory stimuli in the attended hemifield showed significantly faster responses when  
 388 this hemifield was expected than unexpected (psychophysics:  $t(30) = -4.56$ ,  $p < 0.001$ , Cohen's  
 389  $d_{av}$  [95% CI] = -0.40 [-0.59, -0.19]; fMRI:  $t(21) = -5.06$ ,  $p < 0.001$ , Cohen's  $d_{av}$  [95% CI] = -  
 390 0.36 [-0.54, -0.18], Table 1 and left panel in Fig. 1D).

391 For both psychophysics and fMRI experiments, the 2 (attended vs unattended) x 2 (expected vs  
 392 unexpected) repeated measures ANOVA on response times for visual stimuli revealed a  
 393 significant main effect of attention (psychophysics:  $F(1, 30) = 109.88$ ,  $p < 0.001$ ,  $\eta_p^2$  [90% CI] =  
 394 0.79 [0.64, 0.84]; fMRI:  $F(1, 21) = 78.69$ ,  $p < 0.001$ ,  $\eta_p^2$  [90% CI] = 0.79 [0.61, 0.85]).

395 Participants responded faster to visual stimuli in their attended than unattended hemifield.

396 Moreover, a significant crossover interaction between attention and expectation was observed  
 397 (psychophysics:  $F(1, 30) = 41.59$ ,  $p < 0.001$ ,  $\eta_p^2$  [90% CI] = 0.58 [0.36, 0.69]; fMRI:  $F(1, 21) =$   
 398 49.29,  $p < 0.001$ ,  $\eta_p^2$  [90% CI] = 0.70 [0.47, 0.79]). The simple main effects showed that

399 participants responded significantly faster to visual targets in the attended hemifield when this  
 400 hemifield was expected than unexpected (psychophysics:  $t(30) = -5.46$ ,  $p < 0.001$ , Cohen's  $d_{av}$

401 [95% CI] = -0.20 [-0.30, -0.11]; fMRI:  $t(21) = -3.94$ ,  $p = 0.001$ , Cohen's  $d_{av}$  [95% CI] = -0.14 [-

402 0.22, -0.06]) (Table 1 and right panel in Fig. 1D). By contrast, they responded significantly more

403 slowly to visual targets in the unattended hemifield when this hemifield was expected than  
 404 unexpected (psychophysics:  $t(30) = 5.44$ ,  $p < 0.001$ , Cohen's  $d_{av}$  [95% CI] = 0.22 [0.12, 0.32];  
 405 fMRI:  $t(21) = 5.79$ ,  $p < 0.001$ , Cohen's  $d_{av}$  [95% CI] = 0.18 [0.09, 0.26]) (Table 1 and right panel  
 406 in Fig. 1D). Importantly, we observed equivalent response time results in the psychophysics and  
 407 the fMRI experiment. As we discussed in the Materials and Methods section, this crossover  
 408 interaction between attention and expectation can be explained by the profile of general and  
 409 spatially selective response probabilities across conditions. Most prominently, when attention  
 410 and expectation are directed to different hemifields as in block type 2, observers need to inhibit  
 411 responses on a greater proportion of trials, leading to slower response times.

412

413 Table 1 approximately here

414

#### 415 ***fMRI results***

##### 416 *Effects of auditory spatial attention separately for auditory and visual stimuli*

417 We first evaluated the main effect of spatial attention, separately for each stimulus modality.  
 418 For auditory stimuli, auditory spatial attention (i.e., A attended vs unattended auditory stimuli)  
 419 increased activations in bilateral thalami, caudates, hippocampi, left fronto-parietal operculum,  
 420 left putamen and in a motor network encompassing the left central sulcus and the right  
 421 cerebellum. The increased activations for auditory stimuli in motor areas can be explained by the  
 422 motor responses that were given to auditory stimuli only in the attended hemifield. Conversely,  
 423 because visual stimuli required a motor response in both hemifields, no attentional effects were  
 424 observed in the motor network for visual stimuli.

425 Reorienting spatial attention to an auditory stimulus presented in the auditory unattended  
426 hemifield (i.e., A unattended vs attended auditory stimuli) induced activations in a  
427 predominantly right lateralized fronto-parietal system encompassing the bilateral superior frontal  
428 gyri (SFG)/sulci and the right postcentral sulcus extending into the intraparietal sulcus (IPS) and  
429 the inferior parietal lobule (IPL). Significant activation increases were also observed in the right  
430 anterior cingulate cortex/SFG, right middle frontal gyrus (MFG) and bilateral insulae (orange  
431 and yellow on the inflated brain in Fig. 2).

432 Likewise, shifting attention to a visual stimulus in the auditory unattended hemifield (i.e., A  
433 unattended vs attended visual stimuli) increased activations in a more bilateral fronto-parietal  
434 network including bilateral SFG, superior frontal, precentral and postcentral sulci extending into  
435 IPS. We also observed activation increases for unattended visual stimuli in the bilateral anterior  
436 cingulate cortices and right anterior insula (blue and yellow on the inflated brain in Fig. 2). Thus,  
437 even though spatial attention was manipulated selectively in the auditory modality, we observed  
438 similar effects for visual and auditory stimuli when they were presented in the hemifield that was  
439 task-irrelevant in audition.

440 For completeness, no significant main effects of auditory attention for auditory or visual stimuli  
441 were found in primary auditory or visual cortices in our selective ROI analysis. However, we  
442 observed a main effect of auditory attention on visual stimuli at threshold significance ( $p = 0.05$ )  
443 in auditory cortices (Table 4).

444

445 *Effects of auditory spatial attention: commonalities and differences between auditory and visual*  
446 *stimuli.*

447 Next, we investigated the extent to which the neural systems engaged by attention shifts are  
 448 common (i.e., amodal) or distinct (i.e., modality-specific) for auditory and visual stimuli. The  
 449 conjunction analysis over sensory modalities showed increased activations for attention shifts  
 450 (i.e., [A unattended > attended auditory stimuli]  $\cap$  [A unattended > attended visual stimuli]) in  
 451 the bilateral SFG and sulci, right anterior cingulate gyrus, right postcentral sulcus extending into  
 452 IPS, and right anterior insula (Table 2 and Fig. 2).

453 Only the right insula, which was also part of the attentional system that was commonly engaged  
 454 by unattended auditory and visual stimuli, showed a stronger attentional effect for auditory than  
 455 visual stimuli (i.e., interaction: [A unattended > attended auditory stimuli] > [A unattended >  
 456 attended visual stimuli]) (Table 2 and Fig. 2).

457 Table 2 and Fig. 2 show the significant clusters of the conjunction analysis: (i) right postcentral  
 458 sulcus /IPS (encircled in yellow on the inflated SPM template and encircled in black in the axial  
 459 slice) and (ii) right SFG (encircled in yellow on the inflated SPM template and encircled in black  
 460 in the axial slice). (iii) The interaction between attention and stimulus modality in the right insula  
 461 is shown in Table 2 and Fig. 2 (encircled in white on the inflated SPM template and encircled in  
 462 black in the axial slice).

463

464 In summary, in line with our behavioral results, our fMRI analysis suggests that the effect of  
 465 auditory spatial attention generalizes to visual stimuli. Spatial reorienting to both auditory and  
 466 visual stimuli in the hemifield that was task-irrelevant selectively in audition increased  
 467 activations in a widespread right lateralized fronto-parietal system (Shomstein and Yantis, 2006;  
 468 Indovina and Macaluso, 2007; Santangelo et al., 2008; Shulman et al., 2009; Doricchi et al.,  
 469 2010). While the right insula exhibited significantly stronger attentional effects for auditory than

470 visual stimuli, we did not observe attentional effects that were truly selective for stimuli from  
471 either the visual or auditory modality. Collectively, these results suggest that spatial attention and  
472 reorienting rely predominantly on neural systems that are interactively shared across sensory  
473 modalities, even though they may be more strongly engaged by stimuli of the sensory modality  
474 where spatial attention is directly manipulated.

475

476 Table 2 approximately here

477 Figure 2 approximately here

478

479 *Effects of auditory spatial expectation separately for auditory and visual stimuli*

480 Auditory stimuli in the expected relative to unexpected hemifield elicited significantly greater  
481 activation in the bilateral medial prefrontal cortices (i.e., anterior portions of the SFG) and the  
482 bilateral precunei/posterior cingulate gyri (Table 3 and Fig. 3, Summerfield et al., 2006).

483

484 Table 3 approximately here

485 Figure 3 approximately here

486

487 By contrast, auditory stimuli in the unexpected relative to expected hemifield increased  
488 activations in a widespread fronto-parietal system encompassing bilateral SFG/MFG and sulci  
489 and the postcentral/parietal sulci extending into IPL. We also observed activation increases for  
490 unexpected auditory stimuli in the bilateral precunei, anterior insulae, anterior and posterior  
491 cingulate gyri, and in the bilateral planum temporalis and superior temporal gyri (STG) previously  
492 implicated in spatial processing (Griffiths and Warren, 2002; Brunetti et al., 2005; Ahveninen et



493 al., 2006) (Table 3, Fig. 4). Critically, the effects of expectation in regions of the auditory  
494 cortices were not observed for unattended relative to attended stimuli ( $p < 0.05$  uncorrected)  
495 suggesting that they were selective for auditory expectation. Surprisingly, unexpected relative to  
496 expected auditory stimuli increased activations also in the bilateral calcarine cortices (Table 3,  
497 Fig. 4).

498  
499 Our selective ROI analysis also revealed higher activations for unexpected relative to expected  
500 auditory stimuli in primary auditory and visual cortices (indicated by the asterisks in Table 4).  
501 Surprisingly, neither whole brain nor ROI analysis revealed any significant effects of spatial  
502 expectation for visual stimuli.

503

504 Table 4 approximately here

505

506 *Effects of auditory spatial expectation: commonalities and differences between auditory and*  
507 *visual stimuli*

508 Our results suggest that a widespread neural system forms spatial expectations selectively for  
509 stimuli from the auditory modality where signal probability was manipulated. Indeed, this was  
510 confirmed by the significant interaction between expectation and stimulus modality ([A  
511 unexpected > expected auditory stimuli] > [A unexpected > expected visual stimuli]) that was  
512 observed in large parts of the neural system showing auditory expectation effects for auditory  
513 stimuli (see areas on the inflated brain with white outline in Fig. 4). By contrast, the conjunction  
514 analyses over stimulus modality did not reveal any significant effects of auditory expectation that  
515 were common to auditory and visual stimuli.

516

517

Figure 4 approximately here

518

519 *Additive and interactive effects of spatial attention and expectation - separately for auditory and*  
 520 *visual stimuli*

521 Finally, we investigated separately for auditory and visual stimuli whether attention and  
 522 expectation effects engage common neural systems in an additive (i.e., conjunction over  
 523 attention and expectation) or interactive fashion (i.e., interaction between attention and  
 524 expectation).

525 For auditory stimuli, neither the whole brain nor the selective ROI analysis (Table 4) revealed  
 526 any significant interaction between attention and expectation. By contrast, the conjunction  
 527 analysis over attention and expectation revealed activation increases jointly for unattended >  
 528 attended and unexpected > expected (i.e., in an additive fashion) in a predominantly right-  
 529 lateralized fronto-parietal system including the bilateral superior/middle frontal gyri/sulci and the  
 530 right postcentral/intraparietal sulcus extending into right IPL. Further, we observed additive  
 531 effects in the right anterior cingulate gyrus and in the bilateral insulae (Fig. 5).

532

533

Figure 5 approximately here

534

535 For visual stimuli, the whole brain analysis did not reveal any significant additive or interactive  
 536 effects for attention and expectation. The ROI analysis revealed a significant interaction between  
 537 attention and expectation in the primary visual cortex, with greater activations for unexpected  
 538 than expected visual stimuli in the attended hemifield, but lower activation for unexpected than

539 expected visual stimuli in the unattended hemifield (indicated by the asterisk in Table 4). As we  
540 discussed in the Materials and Methods section, this interaction between attention and  
541 expectation may potentially be caused by differences in response probabilities, response times  
542 and associated processes of response inhibition that are increased for block type 2 relative to  
543 block type 1.

544

545

546 *Effect of ‘awareness of auditory expectation manipulation’*

547 Only four out of the 22 participants were not aware of the spatial expectation manipulation in  
548 audition. For completeness, we therefore investigated whether the expectation effects for  
549 auditory stimuli depended on observers’ explicit knowledge about auditory signal probability.  
550 Hence, at the second between participants level we compared the auditory expectation effects  
551 between these four ‘unaware’ and 18 ‘aware’ participants (i.e., interaction between unexpected >  
552 expected for auditory stimuli for ‘aware’ vs ‘unaware’). However, this interaction did not reveal  
553 significant clusters (whole brain corrected). By contrast, a conjunction-null conjunction analysis  
554 over both groups replicated the effects for unexpected relative to expected auditory stimuli in  
555 planum temporale, anterior insula and parietal cortex. These results suggest that explicit  
556 knowledge may not be required for the brain to express activation increases signaling a  
557 prediction error for unexpected auditory stimuli.

558

559

560 **Discussion**

561 The current study was designed to investigate whether the brain allocates attentional resources  
562 and forms expectations over space separately or interactively across the senses. To dissociate the  
563 effects of spatial attention and expectation we orthogonally manipulated spatial attention as  
564 response requirement and expectation as stimulus probability over space selectively in audition  
565 and assessed their effects on neural and behavioral responses in audition and vision.

566 Consistent with previous research, our behavioral results show that participants responded  
567 significantly faster to visual stimuli that were presented in the hemifield where auditory stimuli  
568 were task-relevant (Spence and Driver, 1996, 1997). In other words, directing observers' spatial  
569 attention to one hemifield selectively in audition impacted participants' response speed to  
570 auditory and visual stimuli, suggesting that attentional resources are at least partly shared across  
571 sensory modalities.

572 Likewise, the neural responses to both auditory and visual stimuli depended on auditory spatial  
573 attention. Irrespective of their sensory modality, unattended relative to attended stimuli increased  
574 activations in a widespread right-lateralized dorsal and ventral fronto-parietal system that has  
575 previously been implicated in sustained spatial attention (e.g., Leitão et al., 2015, 2017) and  
576 spatial (re)orienting and contextual updating based on attentional cuing paradigms that conflated  
577 attention and expectation (Nobre et al., 2000; Macaluso et al., 2002; Kincade et al. 2005;  
578 Bressler et al., 2008; Santangelo et al., 2008). By orthogonally manipulating task-relevance (i.e.,  
579 response requirement) and expectation (i.e., signal probability), the current study allowed us to  
580 attribute these fronto-parietal activations to attentional mechanisms. Our results corroborate that  
581 the brain has only limited abilities to define spatial locations as task-relevant or irrelevant  
582 independently for audition and vision (Eimer and Schröger, 1998; Eimer, 1999; Macaluso,  
583 2010). As a result, visual stimuli engaged spatial reorienting even though they should have been

584 attended equally in both hemifields. Conversely, auditory stimuli induced attentional reorienting,  
585 even though they should have been ignored as task-irrelevant. The – at least to some extent –  
586 ‘amodal’ definition of spatial task-relevance may also explain the extensive activations that we  
587 observed for ‘unattended’ stimuli not only in the ventral, but also the dorsal attentional network  
588 that is typically more associated with sustained attention. Greater sustained attention may be  
589 required for stimuli in the ‘auditory unattended’ hemifield, because the brain needs to decide  
590 whether to respond (i.e., visual stimuli) or not to respond (i.e., auditory stimuli; see also Indovina  
591 and Macaluso, 2007; Santangelo et al., 2008 for further discussion). In summary, our behavioral  
592 and neuroimaging findings suggest that spatial attention, when defined as task-relevance,  
593 operates interactively across the senses.

594 Next, we asked whether the neural systems encode spatial signal probability independently  
595 across audition and vision. Behaviorally, we observed faster responses to expected than  
596 unexpected stimuli irrespective of sensory modality in the task-relevant hemifield. Yet,  
597 surprisingly we observed faster responses for unexpected than expected visual stimuli in the  
598 task-irrelevant hemifield (n.b. auditory stimuli did not require a response in the task-irrelevant  
599 hemifield). Hence, we observed a significant interaction between attention and expectation for  
600 visual response times. As discussed in the Materials and Methods section and in greater detail in  
601 Zuanazzi and Noppeney (2018), this interaction for visual response times most likely results  
602 from the differences in the general response probability across conditions. The response  
603 probability is greater when attention and expectation are congruent and directed to the same  
604 hemifield (90% of the trials in blocks of type 1) than when they are directed to different  
605 hemifields (60% of the trials in blocks of type 2) (Fig. 1A and 1B). Put differently, observers

606 need to inhibit their responses on a greater proportion of trials in block type 2, when the  
607 frequency of task-irrelevant auditory stimuli in the auditory unattended hemifield is high.

608

609 Critically, fMRI allows us to move beyond response times and track the neural processes across  
610 the entire visual and auditory processing hierarchy irrespective of whether (e.g., auditory) stimuli  
611 are responded to. This provides us with the opportunity to investigate whether the brain forms  
612 expectations or spatial event probability maps separately for different sensory modalities. Based  
613 on the notion of predictive coding, we would then expect activation increases signaling a  
614 prediction error for stimuli that are presented at unexpected spatial locations (Rao and Ballard,  
615 1999; Friston, 2005). Indeed, spatially unexpected relative to expected auditory stimuli increased  
616 activations as a prediction error signal in the plana temporalia that are critical for auditory spatial  
617 encoding as well as higher order fronto-parietal areas. These results suggest that the planum  
618 temporale forms spatial prediction error signals for spatial unexpected auditory stimuli that then  
619 propagate up the hierarchy into fronto-parietal areas (Friston, 2005). Alternatively, prediction  
620 error signals in the planum temporale may trigger the fronto-parietal attentional system leading  
621 to spatial reorienting (den Ouden et al., 2012). Our design and the sluggishness of the BOLD-  
622 response make it difficult to dissociate between these two explanations for the fronto-parietal  
623 activations. Future EEG/MEG studies may be able to disentangle whether the expression of  
624 prediction error signals in the planum temporale may subsequently trigger attentional reorienting  
625 in the fronto-parietal system.

626 Critically, however, we observed activation increases only for auditory stimuli when presented  
627 in the auditory unexpected hemifield, but not for visual stimuli. In fact, even the visual cortex

628 showed activation increases only for unexpected auditory stimuli potentially mediated via direct  
629 connectivity from auditory areas or top-down modulation from parietal cortices.

630 Likewise, activation increases for spatially expected stimuli were observed selectively for  
631 audition in the medial prefrontal cortex that has previously been implicated in forming  
632 representations consistent with one's expectations (Summerfield et al., 2006). Hence, in line with  
633 the notion of predictive coding, higher order areas such as the medial prefrontal cortex form  
634 representations when stimuli match our spatial expectations, while sensory and potentially  
635 fronto-parietal areas signal a prediction error when our spatial expectations are violated (Rao and  
636 Ballard, 1999; Friston, 2005). Critically, spatial expectations and prediction error signals were  
637 formed in a modality-specific fashion selectively for audition, where stimulus probability was  
638 explicitly manipulated. In fact, we did not observe any significant positive or negative  
639 expectation effects for visual stimuli anywhere in the brain even at an uncorrected threshold of  $p$   
640  $< 0.2$  at the cluster level. These results suggest that the neural systems can form and compute  
641 spatial expectations that encode the probability of stimulus occurrence separately for different  
642 sensory modalities.

643

644 Finally, we asked separately for audition and vision whether spatial attention and expectations  
645 influence these neural responses in an additive or interactive fashion. Recent accounts of  
646 predictive coding suggest that attention may increase the precision of prediction errors  
647 potentially leading to an increase in prediction error signals (i.e., BOLD-response enhancement  
648 for unexpected relative expected stimuli) in the attended hemifield (Feldman and Friston, 2010;  
649 Aukstulewicz and Friston, 2015). However, contrary to this prediction, spatial attention and  
650 expectation did not interact in the auditory modality but influenced neural responses in this

651 system in an additive fashion. Our conjunction analysis over spatial attention and expectation  
 652 revealed a dorsal and ventral fronto-parietal network that was jointly recruited by spatial  
 653 reorienting as well as expectation violations in audition (n.b. which may in turn trigger spatial  
 654 reorienting). By contrast, in primary visual areas we observed a significant interaction between  
 655 spatial attention and expectation selectively for visual stimuli (ROI analysis, Table 4).  
 656 Activations for visual stimuli were greater when attention and expectation were directed to  
 657 different hemifields than to the same hemifield. This activation profile mimics the pattern that we  
 658 observed for behavioral response times and can be found at a lower threshold of significance  
 659 throughout the motor system (e.g., primary motor cortex and cerebellum). It may thus be most  
 660 likely mediated by top-down influences from response selection processes onto sensory cortices  
 661 (van Elk et al., 2010; Gutteling et al., 2013; Gutteling et al., 2015). The interaction between  
 662 attention and expectation in our study highlights processes of expectation (or stimulus  
 663 history/probability) that depend on whether these stimuli were task-relevant (i.e., targets) or  
 664 irrelevant (i.e., non-targets). It thereby contributes to the growing literature that reveals the  
 665 importance of selection history (i.e., the probability of targets vs non-targets or distractors) on  
 666 spatial (and other) priority maps (Awh et al., 2012; Lamy and Kristjánsson, 2013; Chelazzi and  
 667 Santandrea, 2018; Theeuwes, 2018).

668

669 In conclusion, our results demonstrate that spatial attention and expectation engage partly  
 670 overlapping neural systems yet differ in their modality-specificity. Attentional resources were  
 671 controlled interactively across audition and vision within a widespread right-lateralized fronto-  
 672 parietal system. By contrast, spatial expectations and prediction error signals were formed in the



673 planum temporale and fronto-parietal cortices selectively for auditory stimuli where stimulus  
674 probability was explicitly manipulated.

675 Future studies need to investigate the extent to which the modality-specificity of spatial  
676 expectations depends on the statistical structure of the multisensory environment. For instance, in  
677 our experiment auditory and visual signals never occurred together thereby promoting an  
678 encoding of signal probability separately for each sensory modality. We therefore need to assess  
679 the impact of correlations between auditory and visual signals on the encoding of signal  
680 probability. Moreover, given the highly factorial nature of our design, we manipulated signal  
681 probability only in audition and assessed the generalization to vision. The reverse experiment  
682 (i.e., manipulating signal probability in vision) could reveal potential differences in the encoding  
683 and generalization of signal probability between audition and vision (see related discussion about  
684 asymmetric links of attentional resources in Spence and Driver, 1997; Ward et al., 2000).  
685 Because auditory events are typically transient and visual objects permanent, the brain may have  
686 developed different strategies for encoding signal probabilities across the senses. Finally, future  
687 studies may manipulate stimulus probability via probabilistic cues rather than stimulus frequency  
688 to further characterize the neural mechanisms mediating prediction and prediction error signals  
689 (e.g., relationship between expectations/predictions and repetition suppression/priming, Wiggs  
690 and Martin, 1998).

691

692 **Author contributions:** A.Z. and U.N. conceived and designed the experiments. A.Z. performed  
693 the experiments. A.Z. and U.N. analyzed the data. A.Z. and U.N. contributed to the writing of the  
694 manuscript.

695

696 **References**

- 697 Ahveninen J, Jaaskelainen IP, Raij T, Bonmassar G, Devore S, Hamalainen M, Levanen S, Lin  
698 F-HH, Sams M, Shinn-Cunningham BG, Witzel T, Belliveau JW (2006) Task-modulated  
699 “what” and “where” pathways in human auditory cortex. *Proc Natl Acad Sci U S A*  
700 103:14608–14613.
- 701 Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L (2010) Stimulus predictability reduces  
702 responses in primary visual cortex. *J Neurosci* 30:2960–2966.
- 703 Ashburner J, Friston KJ (2005) Unified segmentation. *Neuroimage* 26:839–851.
- 704 Auksztulewicz R, Friston K (2015) Attentional enhancement of auditory mismatch responses: a  
705 DCM/MEG Study. *Cereb Cortex* 25:4273–4283.
- 706 Awh E, Belopolsky A V., Theeuwes J (2012) Top-down versus bottom-up attentional control: a  
707 failed theoretical dichotomy. *Trends Cogn Sci* 16:437–443.
- 708 Blignaut P (2009) Fixation identification: the optimum threshold for a dispersion algorithm.  
709 *Atten Percept Psychophys* 71:881–895.
- 710 Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
- 711 Brefczynski JA, DeYoe EA (1999) A physiological correlate of the “spotlight” of visual  
712 attention. *Nat Neurosci* 2:370–374.
- 713 Bressler DW, Fortenbaugh FC, Robertson LC, Silver MA (2013) Visual spatial attention  
714 enhances the amplitude of positive and negative fMRI responses to visual stimulation in an  
715 eccentricity-dependent manner. *Vision Res* 85:104–112.
- 716 Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M (2008) Top-down control of  
717 human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention.  
718 *JNeurosci* 28:10056–10061.

- 719 Brunetti M, Belardinelli P, Caulo M, Del Gratta C, Della Penna S, Ferretti A, Lucci G, Moretti  
720 A, Pizzella V, Tartaro A, Torquati K, Olivetti Belardinelli M, Romani GL (2005) Human  
721 brain activation during passive listening to sounds from different locations: An fMRI and  
722 MEG study. *Hum Brain Mapp* 26:251–261.
- 723 Carrasco M (2011) Visual attention: The past 25 years. *Vision Res* 51:1484–1525.
- 724 Chelazzi L, Santandrea E (2018) The time constant of attentional control: short, medium and  
725 long (Infinite?). *J Cogn* 1:1–3.
- 726 Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain: from  
727 environment to theory of mind. *Neuron* 58:306–324.
- 728 Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the  
729 brain. *Nat Rev Neurosci* 3:201–215.
- 730 Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain: from  
731 environment to theory of mind. *Neuron* 58:306–324.
- 732 Den Ouden HEM, Kok P, de Lange FP (2012) How prediction errors shape perception, attention,  
733 and motivation. *Front Psychol* 3:1–12.
- 734 Didelot A, Mauguière F, Redouté J, Bouvard S, Lothe A, Reilhac A, Hammers A, Costes N,  
735 Ryvlin P (2010) Voxel-based analysis of asymmetry index maps increases the specificity of  
736 18F-MPPF PET abnormalities for localizing the epileptogenic zone in temporal lobe  
737 epilepsies. *J Nucl Med* 51:1732–1740.
- 738 Doherty JR, Rao A, Mesulam MM, Nobre AC (2005) Synergistic effect of combined temporal  
739 and spatial expectations on visual attention. *J Neurosci* 25:8259–8266.

- 740 Doricchi F, Macci E, Silvetti M, Macaluso E (2010) Neural correlates of the spatial and  
 741 expectancy components of endogenous and stimulus-driven orienting of attention in the  
 742 posner task. *Cereb Cortex* 20:1574–1585.
- 743 Downing CJ (1988) Expectancy and visual-spatial attention: effects on perceptual quality. *J Exp*  
 744 *Psychol Hum Percept Perform* 14:188–202.
- 745 Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new  
 746 SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging  
 747 data. *Neuroimage* 25:1325–1335.
- 748 Eimer M (1999) Can attention be directed to opposite locations in different modalities? An ERP  
 749 study. *Clin Neurophysiol* 110:1252–1259.
- 750 Eimer M, Schröger E (1998) ERP effects of intermodal attention and cross-modal links in spatial  
 751 attention. *Psychophysiology* 35:313–327.
- 752 Feldman H, Friston KJ (2010) Attention, uncertainty, and free-energy. *Front Hum Neurosci* 4:1–  
 753 23.
- 754 Friston K (2005) A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360:815–  
 755 836.
- 756 Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ (1995) Statistical  
 757 parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp* 2:189–  
 758 210.
- 759 Gardner WG, Martin KD (1995) HRTF measurements of a KEMAR. *J. Acoust Soc Am* 97:  
 760 3907–3908.

- 761 Geng JJ, Behrmann M (2002) Probability cuing of target location facilitates visual search  
 762 implicitly in normal participants and patients with hemispatial neglect. *Psychol Sci* 13:520–  
 763 525.
- 764 Geng JJ, Behrmann M (2005) Spatial probability as an attentional cue in visual search. *Percept*  
 765 *Psychophys* 67:1252–1268.
- 766 Griffiths TD, Warren JD (2002) The planum temporale as a computational hub. *Trends Neurosci*  
 767 25:348–353.
- 768 Gutteling TP, Park SY, Kenemans JL, Neggers SFW (2013) TMS of the anterior intraparietal  
 769 area selectively modulates orientation change detection during action preparation. *J*  
 770 *Neurophysiol* 110:33–41.
- 771 Gutteling TP, Petridou N, Dumoulin SO, Harvey BM, Aarnoutse EJ, Kenemans JL, Neggers  
 772 SFW (2015) Action preparation shapes processing in early visual cortex. *J Neurosci*  
 773 35:6472–6480.
- 774 Indovina I, Macaluso E (2007) Dissociation of stimulus relevance and saliency factors during  
 775 shifts of visuospatial attention. *Cereb Cortex* 17:1701–1711.
- 776 Kincade JM, Abrams RA, Astafiev S V., Shulman GL, Corbetta M (2005) An event-related  
 777 functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of  
 778 attention. *J Neurosci* 25:4593–4604.
- 779 Kok P, Jehee JFM, de Lange FP (2012a) Less is more: expectation sharpens representations in  
 780 the primary visual cortex. *Neuron* 75:265–270.
- 781 Kok P, Rahnev D, Jehee JFM, Lau HC, de Lange FP (2012b) Attention reverses the effect of  
 782 prediction in silencing sensory signals. *Cereb Cortex* 22:2197–2206.

- 783 Krumbholz K, Nobis EA, Weatheritt RJ, Fink GR (2009) Executive control of spatial attention  
784 shifts in the auditory compared to the visual modality. *Hum Brain Mapp* 30:1457–1469.
- 785 Lamy DF, Kristjansson A (2013) Is goal-directed attentional guidance just intertrial priming? A  
786 review. *J Vis* 13:1–19.
- 787 Leitão J, Thielscher A, Lee H, Tuennerhoff J, Noppeney U (2017) Transcranial magnetic  
788 stimulation of right inferior parietal cortex causally influences prefrontal activation for  
789 visual detection. *Eur J Neurosci* 46:2807–2816.
- 790 Leitão J, Thielscher A, Tünnerhoff J, Noppeney U (2015) Concurrent TMS-fMRI reveals  
791 interactions between dorsal and ventral attentional systems. *J Neurosci Neurosci* 35:11445–  
792 11457.
- 793 Lipschutz B, Kolinsky R, Damhaut P, Wikler D, Goldman S (2002) Attention-dependent  
794 changes of activation and connectivity in dichotic listening. *Neuroimage* 17:643–656.
- 795 Macaluso E (2010) Orienting of spatial attention and the interplay between the senses. *Cortex*  
796 46:282–297.
- 797 Macaluso E, Frith CD, Driver J (2002) Supramodal effects of covert spatial orienting triggered  
798 by visual or tactile events. *J Cogn Neurosci* 14:389–401.
- 799 Macaluso E, Patria F (2007) Spatial re-orienting of visual attention along the horizontal or the  
800 vertical axis. *Exp Brain Res*:23–34.
- 801 McDonald JJ, Teder-Sälejärvi WA, Hillyard SA (2000) Involuntary orienting to sound improves  
802 visual perception. *Nature* 407:906–908.
- 803 Nobre AC, Gitelman DR, Dias EC, Mesulam MM (2000) Covert visual spatial orienting and  
804 saccades: overlapping neural systems. *Neuroimage* 11:210–216.

- 805 Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory.  
 806 *Neuropsychologia* 9:97–113.
- 807 Posner MI (1980) Orienting of attention. *Q J Exp Psychol* 32:3–25.
- 808 Posner MI, Snyder CRR, Davidson BJ (1980) Attention and the detection of signals. 109:160–  
 809 174.
- 810 Rao RPN, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation  
 811 of some extra-classical receptive-field effects. *Nat Neurosci* 2:79–87.
- 812 Rohe T, Noppeney U (2015) Cortical hierarchies perform Bayesian Causal Inference in  
 813 multisensory perception. *PLOS Biol* 13: e1002073.
- 814 Rohe T, Noppeney U (2016) Distinct computational principles govern multisensory integration  
 815 in primary sensory and association cortices. *Curr Biol* 26:509–514.
- 816 Rohe T, Noppeney U (2018) Reliability-weighted integration of audiovisual signals can be  
 817 modulated by top-down control. *Eneuro* 5:1–20.
- 818 Santangelo V, Macaluso E (2012) Spatial attention and audiovisual processing. In: *The New*  
 819 *Handbook of Multisensory Processing* (Stein BE, ed), pp 359–370. Cambridge, MA: The  
 820 MIT Press.
- 821 Santangelo V, Olivetti Belardinelli M, Spence C, Macaluso E (2008) Interactions between  
 822 voluntary and stimulus-driven spatial attention mechanisms across sensory modalities. *J*  
 823 *Cogn Neurosci* 21:2384–2397.
- 824 Shomstein S, Yantis S (2006) Parietal cortex mediates voluntary control of spatial and nonspatial  
 825 auditory attention. *J Neurosci* 26:435–439.

- 826 Shulman GL, Astafiev SV, Franke D, Pope DLW, Abraham Z, Mcavoy MP, Corbetta M (2009)  
827 Interaction of stimulus-driven reorienting and expectation in ventral and dorsal fronto-  
828 parietal and basal ganglia-cortical networks. *J Neurosci* 29:4392–4407.
- 829 Spence C, Driver J (1996) Audiovisual links in endogenous covert spatial attention. *J Exp*  
830 *Psychol Hum Percept Perform* 22:1005–1030.
- 831 Spence C, Driver J (1997) Audiovisual links in exogenous covert spatial orienting. *Percept*  
832 *Psychophys* 59:1–22.
- 833 Spence C, Pavani F, Driver J (2000) Crossmodal links between vision and touch in covert  
834 endogenous spatial attention. *J Exp Psychol Hum Percept Perform* 26:1298–1319.
- 835 Stekelenburg JJ, Vroomen J (2012) Electrophysiological correlates of predictive coding of  
836 auditory location in the perception of natural audiovisual events. *Front Integr Neurosci* 6:1–  
837 7.
- 838 Summerfield C, Egner T (2009) Expectation (and attention) in visual cognition. *Trends Cogn Sci*  
839 13:403–409.
- 840 Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J (2006) Predictive codes for  
841 forthcoming perception in the frontal cortex. *Science* (80- ) 314:1311–1314.
- 842 Summerfield C, Monti JMP, Trittschuh EH, Mesulam M, Egner T (2008) Neural repetition  
843 suppression reflects fulfilled perceptual expectations. *Nat Neurosci* 11:1004–1006.
- 844 Theeuwes J (2018) Visual selection: usually fast and automatic; seldom slow and volitional. *J*  
845 *Cogn* 1:1–15.
- 846 Thirion B, Pinel P, Mériaux S, Roche A, Dehaene S, Poline J-B (2007) Analysis of a large fMRI  
847 cohort: Statistical and methodological issues for group analyses. *Neuroimage* 35:105–120.



848 Tootell RBH, Hadjikhani N, Hall EK, Marrett S, Vanduffel W, Vaughan JT, Dale AM (1998)  
849 The retinotopy of visual spatial attention. *Neuron* 21:1409–1422.

850 van Elk M, van Schie HT, Neggers SFW, Bekkering H (2010) Neural and temporal dynamics  
851 underlying visual selection for action. *J Neurophysiol* 104:972–983.

852 Ward LM, McDonald JJ, Lin D (2000) On asymmetries in cross-modal spatial attention  
853 orienting. *Percept Psychophys* 62:1258–1264.

854 Wiggs CL, Martin A (1998) Properties and mechanisms of perceptual priming. *Curr Opin*  
855 *Neurobiol* 8:227–233.

856 Wu CT, Weissman DH, Roberts KC, Woldorff MG (2007) The neural circuitry underlying the  
857 executive control of auditory spatial attention. *Brain Res* 1134:187–198.

858 Zuanazzi A, Noppeney U (2018) Additive and interactive effects of spatial attention and  
859 expectation on perceptual decisions. *Sci Rep*:1–12.

860

861

862

863

864

865

866

867

868

869

870

871 **Figure legends**

872

873 **Figure 1.** Experimental design, example stimuli of the psychophysics and fMRI experiment and  
874 behavioral results of the fMRI experiment.

875 **A**, The factorial design manipulated: auditory (A) spatial attention (attended hemifield – full  
876 pattern, vs unattended hemifield – striped pattern), A spatial expectation (expected hemifield –  
877 dark shade, vs unexpected hemifield– light shade) and stimulus modality (auditory modality –  
878 orange, vs visual modality – blue). For illustration purposes and analysis, we pooled over  
879 stimulus locations (left/right). Presence vs absence of response requirement is indicated by the  
880 hand symbol. **B**, Number of auditory (orange) and visual (blue) trials in the 2 (A attended vs  
881 unattended) x 2 (A expected vs unexpected) design. Presence vs absence of response requirement  
882 is indicated by the hand symbol. The fraction of the area indicated by the ‘Response’ hand  
883 symbol pooled over the two bars of one particular block type (e.g., block type 1) represents the  
884 ‘general response probability’ (i.e., the overall probability that a response is required on a  
885 particular trial); the general response probability is greater for block type 1 (90%), where  
886 attention and expectation are congruent, than block type 2 (60%), where they are incongruent.  
887 The fraction of the area indicated by the ‘Response’ hand symbol for each bar represents the  
888 ‘spatially selective response probability’, i.e., the probability that the observer needs to make a  
889 response conditioned on the signal being presented in a particular hemifield; the spatially  
890 selective response probability is greater when unattended signals are presented in the unexpected  
891 (71.4%) than expected (38.4%) hemifield. **C**, fMRI runs included ten blocks of 20 trials  
892 alternating with fixation periods. A fixation cross was presented throughout the entire run. Its  
893 color indicated: white = fixation period; green or pink = activation period with auditory attention

894 directed to the left (or right) hemifield. On each trial participants were presented with an auditory  
 895 or visual stimulus (100 ms duration) either in their left or right hemifield. They were instructed  
 896 to respond as fast and accurately as possible with their right index finger within a response  
 897 window of 1400 ms. **D**, Bar plots show response times (across subjects' mean ( $\pm$ SEM)) for each  
 898 of the six conditions with response requirements in the fMRI experiment. The brackets and stars  
 899 indicate significance of main effects and interactions. \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ . Audition:  
 900 orange; vision: blue; attended: full pattern; unattended: striped pattern; expected: dark shade;  
 901 unexpected: light shade.

902  
 903  
 904 **Figure 2.** Auditory (A) unattended > attended for auditory and visual stimuli.  
 905 Activation increases for A unattended > attended stimuli for auditory (AUD, orange, height  
 906 threshold:  $p < 0.001$ , uncorr., extent threshold  $k > 0$  voxels) and visual (VIS, blue, height  
 907 threshold:  $p < 0.001$ , uncorr., extent threshold  $k > 0$  voxels) stimuli (overlap: yellow) are  
 908 rendered on an inflated canonical brain. The conjunction of A unattended > attended for auditory  
 909 and visual stimuli is encircled in yellow (height threshold:  $p < 0.001$ , uncorr., extent threshold  $k$   
 910  $> 0$  voxels). Activation increases for A unattended > attended that are greater for auditory than  
 911 visual stimuli (i.e., interaction) are encircled in white (height threshold:  $p < 0.001$ , uncorr., extent  
 912 threshold  $k > 0$  voxels). Bar plots show the parameter estimates (across participants mean  $\pm$   
 913 SEM, averaged across all voxels in the black encircled cluster) in the (i) right  
 914 postcentral/intraparietal sulcus, (ii) right superior frontal gyrus and (iii) right anterior insula that  
 915 are displayed on axial slices of a mean image created by averaging the subjects' normalized  
 916 structural images. The bar graphs represent the size of the effect pertaining to BOLD magnitude

917 in nondimensional unit (corresponding to percentage whole-brain mean). Audition: orange;  
 918 vision: blue; attended: full pattern; unattended: striped pattern; expected: dark shade;  
 919 unexpected: light shade.

920

921

922 **Figure 3.** Auditory (A) expected > unexpected for auditory and visual stimuli.

923 Activation increases for A expected > unexpected auditory stimuli (orange) are rendered on an  
 924 inflated canonical brain; they are encircled in white if they are significantly greater for auditory  
 925 than visual stimuli (i.e., interaction). Height threshold of  $p < 0.001$ , uncorrected; extent threshold  
 926  $k > 0$  voxels.

927 Bar plots show the parameter estimates (across participants mean  $\pm$  SEM, averaged across all  
 928 voxels in the black encircled cluster) in the medial prefrontal cortices (i.e., anterior portions of  
 929 the superior frontal gyri) that are displayed on axial slices of a mean image created by averaging  
 930 the subjects' normalized structural images; the bar graphs represent the size of the effect in  
 931 nondimensional unit (corresponding to percentage whole-brain mean). Audition: orange; vision:  
 932 blue; attended: full pattern; unattended: striped pattern; expected: dark shade; unexpected: light  
 933 shade.

934

935

936 **Figure 4.** Auditory (A) unexpected > expected for auditory and visual stimuli.

937 Activation increases for A unexpected > expected stimuli for auditory stimuli (orange) are  
 938 rendered on an inflated canonical brain; they are encircled in white if they are significantly

939 greater for auditory than visual stimuli (i.e., interaction). Height threshold of  $p < 0.001$ ,  
 940 uncorrected; extent threshold  $k > 0$  voxels.  
 941 Bar plots show the parameter estimates (across participants mean  $\pm$  SEM, averaged across all  
 942 voxels in the black encircled cluster) in (i) bilateral superior temporal gyri and bilateral  
 943 intraparietal sulci, (ii) bilateral superior frontal gyri and (iii) bilateral calcarine cortices that are  
 944 displayed on axial slices of a mean image created by averaging the subjects' normalized  
 945 structural images. The bar graphs represent the size of the effect in nondimensional unit  
 946 (corresponding to percentage whole-brain mean). Audition: orange; vision: blue; attended: full  
 947 pattern; unattended: striped pattern; expected: dark shade; unexpected: light shade.

948  
 949  
 950 **Figure 5.** Additive effects of auditory (A) attention and expectation in audition (AUD).  
 951 Activation increases common (i.e., conjunction) for A attention and expectation main effects in  
 952 the auditory modality ( $[A \text{ unattended} > \text{attended AUD}] \cap [A \text{ unexpected} > \text{expected AUD}]$ ) are  
 953 rendered in orange on an inflated canonical brain; height threshold of  $p < 0.001$ , uncorrected;  
 954 extent threshold  $k > 0$  voxels.

962 **Tables**963 **Table 1.** Behavioral results.

964 Group mean reaction times (RT) for each stimulus modality in each condition for the  
 965 psychophysics and fMRI experiments. Standard errors of the mean (SEM) are given in  
 966 parentheses.

Experiment	Auditory modality		Visual modality			
	+att +exp	+att -exp	+att +exp	+att -exp	-att +exp	-att -exp
Psychophysics						
RT (ms) (SEM)	530.7 (17.1)	566.8 (15.3)	446.9 (10.2)	458.3 (9.4)	487 (11.4)	472.6 (11.8)
fMRI						
RT (ms) (SEM)	508.4 (24.5)	552.9 (27.4)	432.3 (13.9)	441.2 (12.9)	467.1 (14.9)	454.2 (14.7)

985 **Table 2.** ‘Amodal’ and modality dependent mechanisms of auditory spatial attention.

986  $p$ -values are FWE corrected at the cluster level for multiple comparisons within the entire brain.

987 Auxiliary uncorrected voxel threshold of  $p < 0.001$ . L, Left; R, right; A, auditory.

Brain regions	MNI coordinates (mm)			z-score (peak)	Cluster size (voxels)	$p_{FWE}$ -value (cluster)
	$x$	$y$	$z$			
[A unattended > attended auditory stimuli] $\cap$ [A unattended > attended visual stimuli]						
R superior frontal gyrus	18	-4	64	4.09	731	0.000
R superior frontal sulcus	28	-6	46	4.38		
R anterior cingulate gyrus	10	18	36	3.67		
L superior frontal gyrus	-14	-10	64	3.69	268	0.005
L superior frontal sulcus	-30	-8	48	4.07		
R postcentral sulcus/ R intraparietal sulcus	42	-32	40	3.74	304	0.003
R anterior insula	30	20	6	4.41	185	0.027
[A unattended > attended auditory stimuli] > [A unattended > attended visual stimuli]						
R anterior insula	38	16	-12	4.32	209	0.016

988

989

990

991

992

993

994

995

996

997

998 **Table 3.** Main effects of auditory spatial expectation for auditory stimuli.

999  $p$ -values are FWE corrected at the cluster level for multiple comparisons within the entire brain.

1000 Auxiliary uncorrected voxel threshold of  $p < 0.001$ . L, Left; R, right; A, auditory.

Brain regions	MNI coordinates (mm)			z-score (peak)	Cluster size (voxels)	$p_{FWE}$ -value (cluster)
	$x$	$y$	$z$			
A Expected > unexpected auditory stimuli						
R superior frontal gyrus	8	54	18	4.91	1458	0.000
L superior frontal gyrus	-6	54	36	5.80		
R precuneus	-4	-52	26	3.90	260	0.006
L precuneus	6	-56	26	3.28		
A Unexpected > expected auditory stimuli						
R superior temporal gyrus	60	-44	16	7.47	18305	0.000
L superior temporal gyrus	-62	-34	14	5.44		
R postcentral sulcus/ R intraparietal sulcus	34	-58	46	5.93		
L postcentral sulcus/ L intraparietal sulcus	-38	-46	42	5.79		
R precuneus	4	-54	54	6.49		
L precuneus	-8	-54	54	6.48		
R anterior insula	38	16	2	7.48		
L anterior insula	-32	16	2	6.94		
R posterior cingulate gyrus/ L posterior cingulate gyrus	4	-28	26	5.18	339	0.001
R anterior cingulate gyrus	8	22	32	5.40	4222	0.000
R superior frontal gyrus	18	2	66	4.14		
L superior frontal gyrus	-26	-8	70	4.23		
L precentral sulcus	-38	0	38	5.08		
R precentral sulcus	40	6	30	4.97	2186	0.000
R middle frontal gyrus	40	34	36	4.42		
L middle frontal gyrus	-34	46	24	4.56	810	0.000
R calcarine cortex	12	-84	8	3.75	680	0.000
L calcarine cortex	-12	-84	6	3.71		

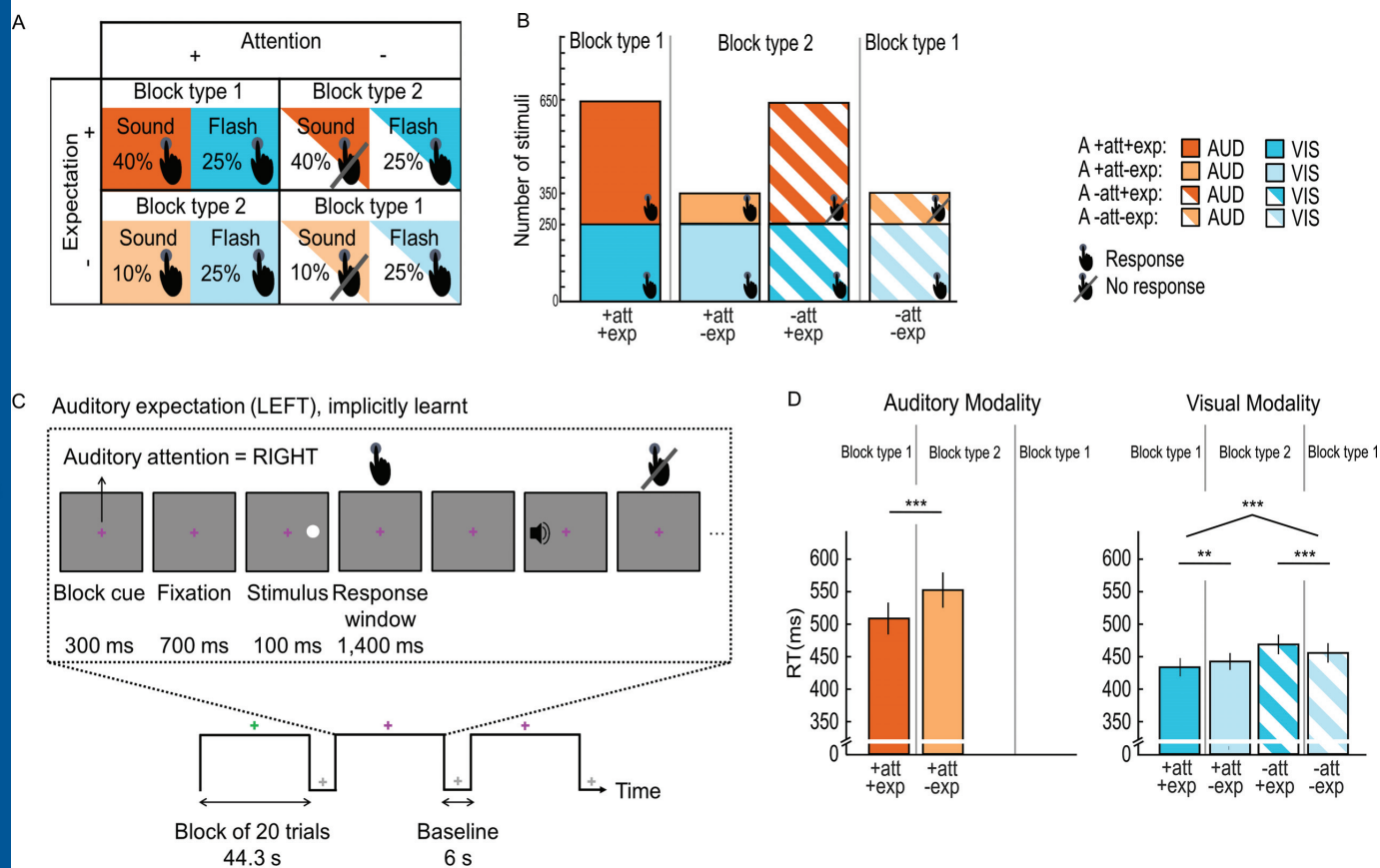


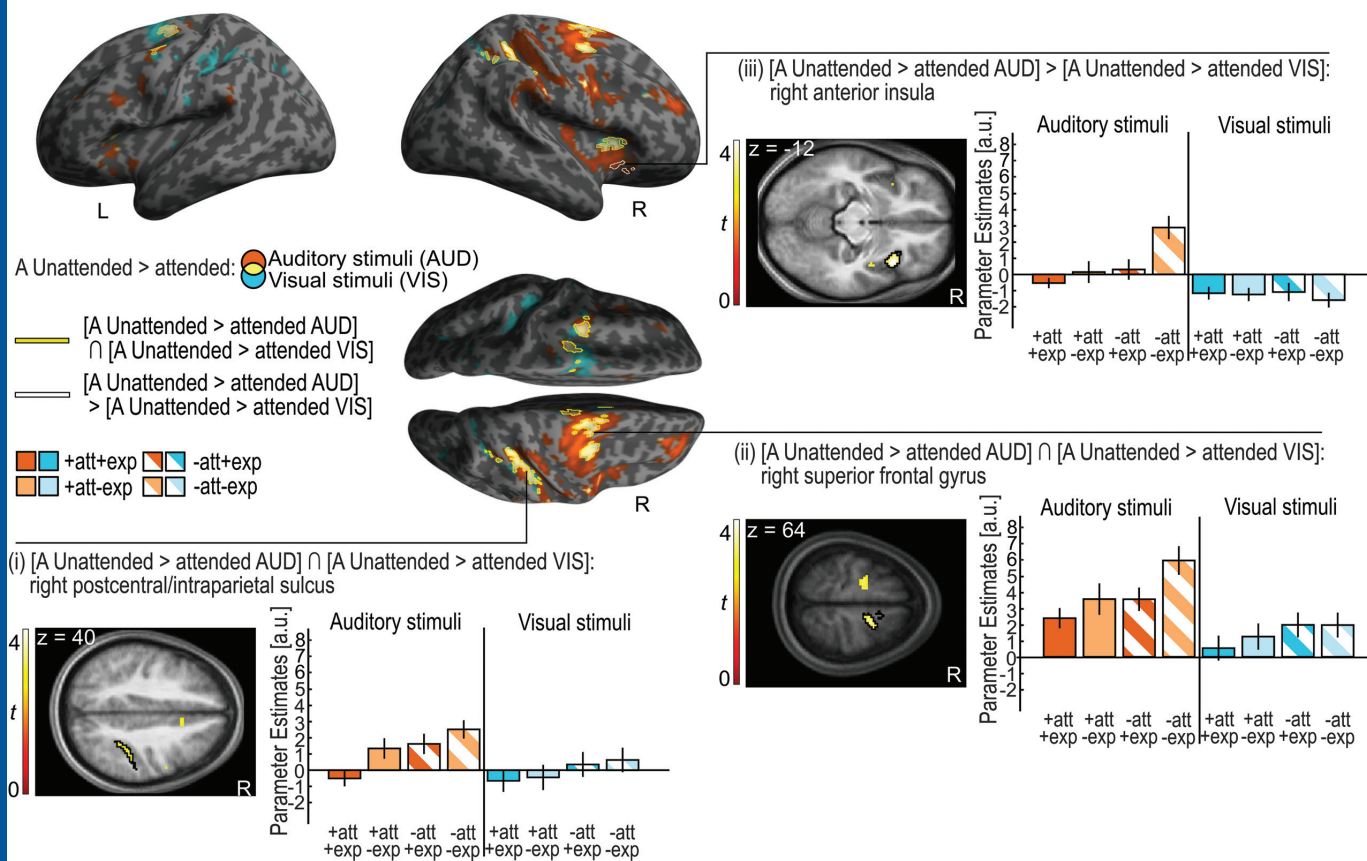
1001 **Table 4.** Results of the ROI analysis for each stimulus modality.  
 1002 ROIs: primary auditory and primary visual cortex. 90% CI of  $\eta_p^2$  is given in square brackets. \*\*  
 1003  $p < 0.01$ , \*  $p < 0.05$ ; A, auditory.

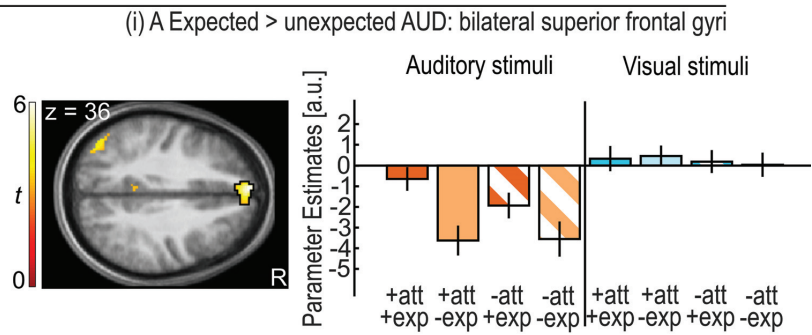
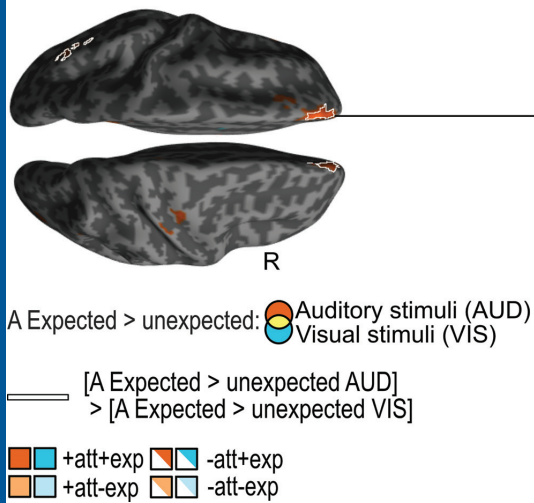
ROIs		Main effect of A attention	Main effect of A expectation	Interaction A attention x expectation
Primary auditory cortex	Auditory stimuli			
	$F(1,21)$	0.148	12.671	0.846
	$p$	0.704	0.002**	0.368
	$\eta_p^2$ [90% CI]	0.007 [0, 0.140]	0.376 [0.106, 0.558]	0.039 [0, 0.223]
	Visual Stimuli			
	$F(1,21)$	4.310	0.213	0.117
Primary visual cortex	$p$	0.050	0.649	0.736
	$\eta_p^2$ [90% CI]	0.170 [0, 0.383]	0.010 [0, 0.154]	0.006 [0, 0.131]
	Auditory stimuli			
	$F(1,21)$	0.165	7.213	2.096
	$p$	0.689	0.014*	0.162
	$\eta_p^2$ [90% CI]	0.008 [0, 0.144]	0.256 [0.032, 0.461]	0.091 [0, 0.296]
	Visual Stimuli			
	$F(1,21)$	0.995	0.054	5.062
	$p$	0.330	0.819	0.035*
	$\eta_p^2$ [90% CI]	0.045 [0, 0.233]	0.003 [0, 0.091]	0.194 [0.008, 0.406]

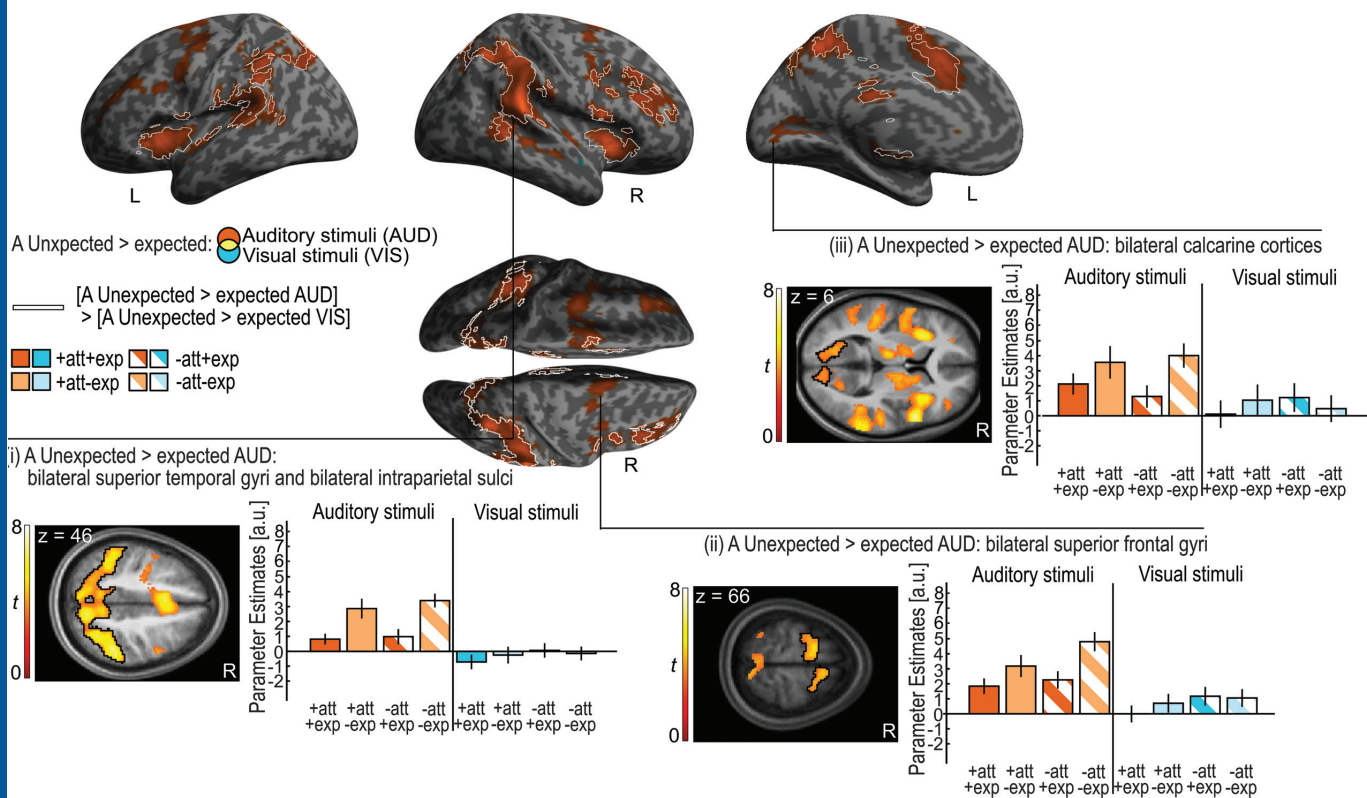
1004

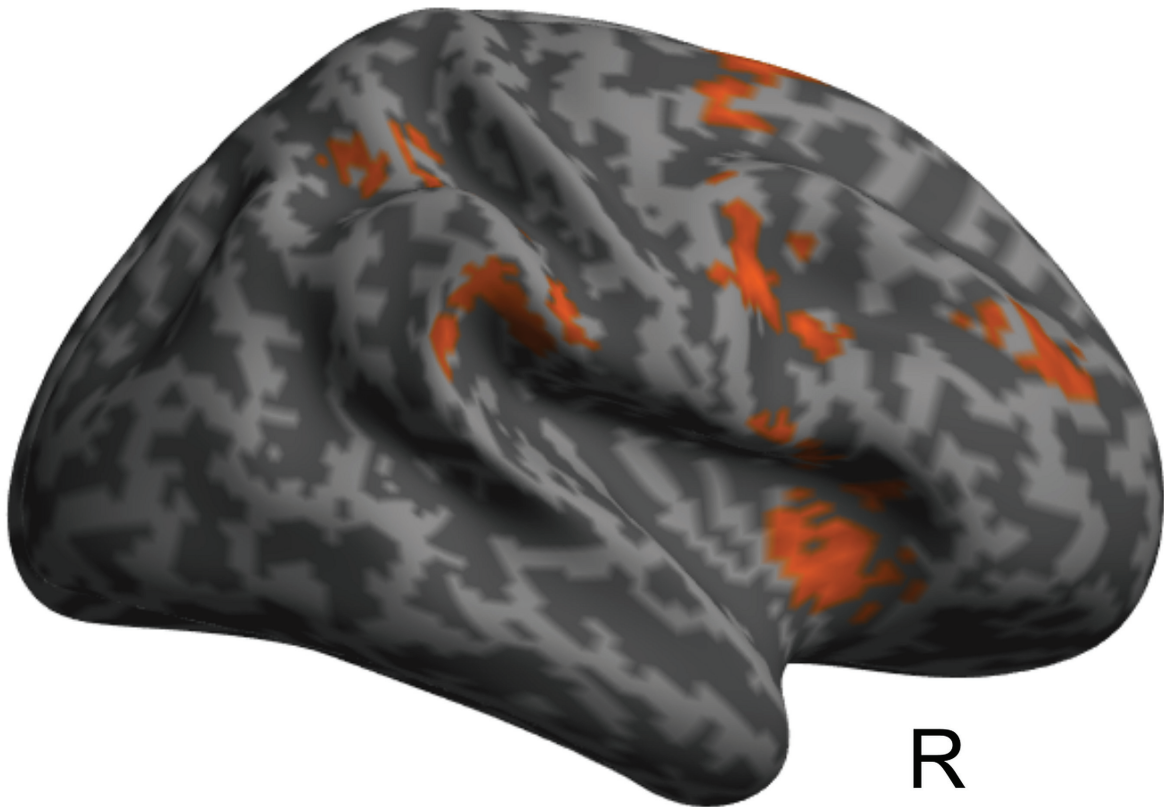
1005











- $[A \text{ unattended} > \text{attended AUD}]$   
 $\cap [A \text{ unexpected} > \text{expected AUD}]$